

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Hillman, Jennifer L.
Corley, Neil C.
Baughn, Mariah R.
- (ii) TITLE OF INVENTION: DELTA 1-PYRROLINE-5-CARBOXYLATE REDUCTASE
HOMOLOG
- (iii) NUMBER OF SEQUENCES: 3
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
 - (B) STREET: 3174 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: Windows
 - (D) SOFTWARE: FastSEQ for Windows Version 2.0b
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: TO BE ASSIGNED
 - (B) FILING DATE: HEREWITH
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Cerrone, Michael C
 - (B) REGISTRATION NUMBER: 39,132
 - (C) REFERENCE/DOCKET NUMBER: PF-0532 US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 650-855-0555
 - (B) TELEFAX: 650-855-0572
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 314 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (vii) IMMEDIATE SOURCE:
 - (A) LIBRARY: PROSNON01

(B) CLONE: 2278458

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Met	Ser	Val	Gly	Phe	Ile	Gly	Ala	Gly	Gln	Leu	Ala	Tyr	Arg	Phe	Thr	1	5	10	15
Ala	Ala	Gly	Ile	Leu	Ser	Ala	His	Lys	Ile	Ile	Ala	Ser	Ser	Pro	Glu	20	25	30	
Met	Asn	Leu	Pro	Thr	Val	Ser	Ala	Leu	Arg	Lys	Met	Gly	Val	Asn	Leu	35	40	45	
Thr	Arg	Ser	Asn	Lys	Glu	Thr	Val	Lys	His	Ser	Asp	Val	Leu	Phe	Leu	50	55	60	
Ala	Val	Lys	Pro	His	Ile	Ile	Pro	Phe	Ile	Leu	Asp	Glu	Ile	Gly	Ala	65	70	75	80
Asp	Val	Gln	Ala	Arg	His	Ile	Val	Val	Ser	Cys	Ala	Ala	Gly	Val	Thr	85	90	95	
Ile	Ser	Ser	Val	Glu	Lys	Lys	Leu	Met	Ala	Phe	Gln	Pro	Ala	Pro	Lys	100	105	110	
Val	Ile	Arg	Cys	Met	Thr	Asn	Thr	Pro	Val	Val	Val	Gln	Glu	Gly	Ala	115	120	125	
Thr	Val	Tyr	Ala	Thr	Gly	Thr	His	Ala	Leu	Val	Glu	Asp	Gly	Gln	Leu	130	135	140	
Leu	Glu	Gln	Leu	Met	Ser	Ser	Val	Gly	Phe	Cys	Thr	Glu	Val	Glu	Glu	145	150	155	160
Asp	Leu	Ile	Asp	Ala	Val	Thr	Gly	Leu	Ser	Gly	Ser	Gly	Pro	Ala	Tyr	165	170	175	
Ala	Phe	Met	Ala	Leu	Asp	Ala	Asp	Gly	Gly	Val	Lys	Met	Gly	Leu	Pro	180	185	190	
Arg	Arg	Leu	Ala	Ile	Gln	Leu	Gly	Ala	Gln	Ala	Leu	Leu	Gly	Ala	Ala	195	200	205	
Lys	Met	Leu	Leu	Asp	Ser	Glu	Gln	His	Pro	Cys	Gln	Leu	Lys	Asp	Asn	210	215	220	
Val	Cys	Ser	Pro	Gly	Gly	Ala	Thr	Ile	His	Ala	Leu	His	Phe	Leu	Glu	225	230	235	240
Ser	Gly	Gly	Phe	Arg	Ser	Leu	Leu	Ile	Asn	Ala	Val	Glu	Ala	Ser	Cys	245	250	255	
Ile	Arg	Thr	Arg	Glu	Leu	Gln	Ser	Met	Ala	Asp	Gln	Glu	Lys	Ile	Ser	260	265	270	
Pro	Ala	Ala	Leu	Lys	Lys	Thr	Leu	Leu	Asp	Arg	Val	Lys	Leu	Glu	Ser	275	280	285	
Pro	Thr	Val	Ser	Thr	Leu	Thr	Pro	Ser	Ser	Pro	Gly	Lys	Leu	Leu	Thr	290	295	300	
Arg	Ser	Leu	Ala	Leu	Gly	Gly	Lys	Lys	Asp							305	310		

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1742 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: PROSNON01
 (B) CLONE: 2278458

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCGCCATCAG	CCGCCCGGGA	GATATCCGCC	GGGGGAGAAT	AGGGTTGCAC	CATCCAGAA	60
GCTGCTGTTA	GCTCGCCGGT	CCTCGGCACG	CGGCCCGTTC	GCCCTGCGC	TGTCGCCCTT	120

TCCCTTAGCG TTAATTCCGG TCCCTCGCTG AGGGGGTTCG TCGGCTCC AGGAGGCGTG 180
AACCGCGGAC CATGAGCGTG GGCTTCATCG GGCCCGGCCA GCTGGCCTAT GCTCTGGCGC 240
GGGGCTTCAC GGCCCGAGGC ATCTGTGCGG CTCACAAGAT AATAGCCAGC TCCCAGAAAA 300
TGAACCTGCC CACGGTGTCC GCGCTCAGGA AGATGGGTGT GAACCTGACA CGCAGCAACA 360
AGGAGACGCT GAAGCACAGC GACGTCCTGT TTCTGGCTGT GAAGCCACAT ATCATCCCTT 420
TCATCTCGTA TGAGATTGGG GCCGACGTGC AAGCCAGACA CATCTGGTCT TCTGTGCGG 480
CTGTTCTCAC CATCACTCTT GTGGAGAAGA AGCTGATGGC ATTCCAGCCA GCCCCCAAAG 540
TGATTTCGCT CATGACCAAC ACACCTGTGG TAGTGCAGGA AGGCGCTACA GTGTACGCCA 600
CGGGCACCCA TGCCCTGGTG GAGGATGGGC AGCTCTTGA GACAGCTCATG ACGCAAGCTGG 660
GCTTCTGCAC TGAGGTGGAA GAGGACCTCA TCGATGCCGT CACGGGGGCTC AGTGGCAGCG 720
GGCCTGCCTA TGCAATTCATG GCTCTGGAGC CATTGGCTGA TGGTGGGGTG AAGATGGGTT 780
TGCCACGGCG CCTGGCAATC CAACTCGGGG CCCAGGCTTT GCTGGGAGCT GCCAAGATGC 840
TGCTGGACTC GGAGCAGCAT CCATGCCAGC TTAAGGACAA GTCTGTCTCC CTTGGGGGAG 900
CCACCATCCA GCGCCTGCAC TTCTTAGAGA GTGGGGGCTT CCGCTCTCTG CTCATCAATG 960
CAGTTGAGCG CTCTCTGTATC CGAACACGAG AGCTACAGTC CATGGCCGAC CAAGAAAAAG 1020
TCTCCCACTG TGCCCTTAAG AAGACCTCTT TAGACAGAGT GAAGCTGGAA TCCCCCACAG 1080
TCTCCACACT GACCCCTCC AGCCCAAGGA AGCTCCTCAC AAGAAGCCTG GCCCTGGGAG 1140
GCAAGAAAGT CTAAGAGTCC ATCTGTCCCT TCTGTGATTC AGAGCCCTTA GTTGAGAGCC 1200
CCTGCCGCCC CTGCCACCCC CTTGCCCCGC TCCCACCAAT GCCCTCTCTC AGCTGTGCAA 1260
GGAGAAAGCA TGCTTAGGAA GTTTTCAGGT CTTGTGATA AAACCTCCTT AAATCTGTTT 1320
AGACCAAGCA ATGCGAGCTT CCTCTCCTGT CCCATGTTGG AAGTTGCTCT GAAGGGGTGG 1380
TAGATGCTGG AAGCCAGACA CAACCTGCG TACGCTGCTC AGTTGGTGA GACTGGGGCT 1440
GGGACTGGAG TCAGCCAGC TGGGAGGAGG GGCTGGGGAG GATCTGCAGC TGAAGCCCCA 1500
GCCAGGGTTG GTGTGATGCC AAGGCAAGT GGTGAGGAGA AAACAGGAAA CGGCGCTTTCT 1560
CTGAATTTGT AAATGGGAAA GAAGTGAGCA ACTTAAGATT GTCACAAATA ATCACAAGT 1620
TACAGATTA GACTGGGTTT ATATTTAACT CTGTCTTAT AGGTACACCA TTTAAAGAGT 1680
GTTATTTAAT GCTAAGTTTA ACTGCTTTAA TAAAGTTTAT TTTTAAATAT CAAAAAATA 1740
AA

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GENBANK
- (B) CLONE: 189498

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Ser	Val	Gly	Phe	Ile	Gly	Ala	Gly	Gln	Leu	Ala	Phe	Ala	Lys	Gly
1			5						10				15		
Phe	Thr	Ala	Ala	Gly	Val	Leu	Ala	Ala	His	Lys	Ile	Met	Ala	Ser	Ser
			20					25					30		
Pro	Asp	Met	Met	Asp	Leu	Ala	Thr	Val	Ser	Ala	Leu	Arg	Lys	Met	Gly
			35				40					45			Val
Lys	Leu	Thr	Pro	His	Asn	Lys	Glu	Thr	Val	Gln	His	Ser	Asp	Val	Leu
			50			55				60					
Phe	Leu	Ala	Val	Lys	Pro	His	Ile	Ile	Pro	Phe	Ile	Leu	Asp	Glu	Ile
65					70				75					80	
Gly	Ala	Asp	Ile	Glu	Asp	Arg	His	Ile	Val	Val	Ser	Cys	Ala	Ala	Gly
			85				90					95			
Val	Thr	Ile	Ser	Ser	Ile	Glu	Lys	Lys	Leu	Ser	Ala	Phe	Arg	Pro	Ala
			100				105					110			
Pro	Arg	Val	Ile	Arg	Cys	Met	Thr	Asn	Thr	Pro	Val	Val	Val	Arg	Glu
			115			120						125			
Gly	Ala	Thr	Val	Tyr	Ala	Thr	Gly	Thr	His	Ala	Gln	Val	Glu	Asp	Gly
			130			135						140			

Arg	Leu	Met	Glu	Gln	Leu	Leu	Ser	Thr	Val	Gly	Phe	Cys	Thr	Glu	Val
145					150					155					160
Glu	Glu	Asp	Leu	Ile	Asp	Ala	Val	Thr	Gly	Leu	Ser	Gly	Ser	Gly	Pro
				165					170					175	
Ala	Tyr	Ala	Phe	Thr	Ala	Leu	Asp	Ala	Asp	Gly	Gly	Val	Lys	Met	Gly
			180					185					190		
Leu	Pro	Arg	Arg	Leu	Ala	Val	Arg	Leu	Gly	Ala	Gln	Ala	Leu	Leu	Gly
		195					200					205			
Ala	Ala	Lys	Met	Leu	Leu	His	Ser	Glu	Gln	His	Pro	Gly	Gln	Leu	Lys
		210				215					220				
Asp	Asn	Val	Ser	Ser	Pro	Gly	Gly	Ala	Thr	Ile	His	Ala	Leu	His	Val
					230					235					240
Leu	Glu	Ser	Gly	Gly	Phe	Arg	Ser	Leu	Leu	Ile	Asn	Ala	Val	Glu	Ala
				245					250					255	
Ser	Cys	Ile	Arg	Thr	Arg	Glu	Leu	Gln	Ser	Met	Ala	Asp	Gln	Glu	Gln
			260					265					270		
Val	Ser	Pro	Ala	Ala	Ile	Lys	Lys	Thr	Ile	Leu	Asp	Lys	Val	Lys	Leu
			275				280					285			
Asp	Ser	Pro	Ala	Gly	Thr	Ala	Leu	Ser	Pro	Ser	Gly	His	Thr	Lys	Leu
		290				295					300				
Leu	Pro	Arg	Ser	Leu	Ala	Pro	Ala	Gly	Lys	Asp					
305					310					315					